

PPT treating viral infections by the picornaviridae class, in chromosomal mapping, tissue typing, forensic biology, or viral serotyping -
PPT XX
PPS XX
XX
Claim 1: Fig 11; 117pp; English.

The present sequence is the protein sequence of novel CM985 perturbagen. To identify molecules that alter a cell's ability to resist or deter rhinovirus (RV) infection, a random primed library of 12×10^6 power 6 clones was constructed from cDNA isolated from placental tissue. This genetic library was transfected into HeLa cells that were previously shown to be susceptible to RV infection. The library-containing population was expanded 9-fold and then exposed to human RV-14 (ATCC VR 284) to identify perturbagens that inhibited the viral life cycle. Clone CM985 was selected. The invention further provides host cells, vectors and gene therapy vectors comprising polynucleotides encoding CM985. The host cells provide for methods for producing polypeptides having viral-related activity, which in turn can be used to identify potential therapeutics. The invention also provides methods for identifying a cellular target that interacts with the perturbagen, e.g. using a yeast two hybrid assay, and for using such targets to screen for putative viral therapeutics. Perturbagen polynucleotides, polypeptides or antibodies may also be used in detection assays (e.g. chromosomal mapping, tissue typing, forensic biology, or viral serotyping), and in predictive medicine.

PT analysing gene expression in human adult liver -
XX
PS Claim 27; SEQ ID No 27475; 658pp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENPs) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver comprising one of 11109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinemia, hyperlipidemia and hypercholesterolemia which
CC is associated with coronary heart disease. ABG47348-ABG5930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp://wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 78 AA:
XX

Query Match 48.9%; Score 135; DB 22; Length 78;
 Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3

Qy	1	DNDGDDNDGGGENG-GRGDNYATDD--MAMMVIMLM-MVYVRLMVYVIMLIM	52
Db	12	DDDDDDDDADNGDGDGGDDVTMMVIMMAMMVIMMVYVIMVMM	67

RESULT 3
ABB28821 ID ABB28821 standard; Peptide: 78 AA.

AC ABB28821;
XX DT 01-FEB-2002 (first entry)
YY

DE Peptide #1472 encoded by breast cell single exon nucleic acid probe.
 XX Human; microarray; single exon probe; gene expression; breast;
 KW disease; cancer.

XX Homo sapiens.
XX WO200157271-A2.
PN

PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00662.
PF YY

PR 04-FEB-2000; 2000US-0180312.
PR 20-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0612366.

PR 21-SEP-2000; 2000US-023687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI, 2001, 100000/54

XX New spatially-addressable set of single exon nucleic acid probes, useful for measuring gene expression in sample derived from human breast. comprises number of single exon nucleic acid probes -

XX Claim 27; SEQ ID NO 11789; 327pp + sequence listing; English.
PS
XX
CC The invention relates to a spatially-addressable set of single exon

CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognostic breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 78 AA;

Query Match 48.9%; Score 135; DB 22; Length 78;

Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 Matches 30; Conservative 13; Indels 4; Gaps 3;

Qy 1 DNDGGDDNDGGGENNG-GDGDNYATDD-MMMYIMMM-MMVYRIMMMVIMILM 52
 Db 12 DDDDGDDDDDDADNDGDDGDDGDDVMMVMMVMMVMMVMM 67

RESULT 4
 ABB34006 standard; Peptide; 78 AA.

XX ABB34006;

XX DT 04-FEB-2002 (first entry)

XX Peptide #l512 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX Homo sapiens.
 XX WO200157277-A2.

XX PD 09-AUG-2001.

XX PP 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human
 hearts -

XX PS Claim 15; SEQ ID No 21217; 530pp; English.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 26641; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 78 AA;

Query Match 48.9%; Score 135; DB 22; Length 78;

Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 Matches 30; Conservative 13; Indels 4; Gaps 3;

Qy 1 DNDGGDDNDGGGENNG-GDGDNYATDD-MMMYIMMM-MMVYRIMMMVIMILM 52
 Db 12 DDDDGDDDDDDADNDGDDGDDGDDVMMVMMVMMVMM 67

RESULT 5
 ABB19447

ID ABB19447 standard; Protein; 78 AA.

XX AC ABB19447;

XX DT 23-JAN-2002 (first entry)

XX DE Protein #1446 encoded by probe for measuring heart cell gene expression.
 XX Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.

XX OS Homo sapiens.

XX PN WO200157274-A2.

XX PD 09-AUG-2001.

XX PP 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human
 hearts -

XX PS Claim 15; SEQ ID No 21217; 530pp; English.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 26641; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 78 AA;

Query Match 48.9%; Score 135; DB 22; Length 78;

Best Local Similarity Matches	53.6%; 30; Conservative	Pred. No. 13;	3.2e-09; Mismatches	9;	Indels	4;	Gaps	3;	
Qy	1 DNDGGDDNDGGENNG-GDGDNYATDD - M-MMMYIMM-MVVRIMVVIVIIMLIM 52								XX DT 06-NOV-2001 (first entry)
Db	12 DDDDDDDDDADNDGDDGDDDTMVMMVMMVMMVMMVMMVMM 67								XX Human bone marrow expressed probe encoded protein SEQ ID NO: 27467.
	RESULT 6								XX Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukaemia; lymphoma; myeloma.
	AAM54774	standard; Protein;	78 AA.						XX Homo sapiens.
	XX	PN	WO200157276-A2.						XX OS
	AC	PD	09-AUG-2001.						XX PN
	XX	PD	09-AUG-2001.						XX XX
	DT	XX	30-JAN-2001; 2001WO-US00668.						XX PR
	XX	XX	04-FEB-2000; 2000US-0180312.						PR 04-FEB-2000; 2000US-0180312.
	DE	XX	26-MAY-2000; 2000US-0207456.						PR 26-MAY-2000; 2000US-0207456.
	DE	XX	30-JUN-2000; 2000US-0608408.						PR 30-JUN-2000; 2000US-0608408.
	DE	XX	03-AUG-2000; 2000US-0632366.						PR 03-AUG-2000; 2000US-0632366.
	DE	XX	21-SEP-2000; 2000US-0234687.						PR 21-SEP-2000; 2000US-0234687.
	OS	XX	27-SEP-2000; 2000US-0236159.						PR 27-SEP-2000; 2000US-0236159.
	OS	XX	04-OCT-2000; 2000GB-02423.						PR 04-OCT-2000; 2000GB-02423.
	XX	PA	(MOLE-) MOLECULAR DYNAMICS INC.						XX XX
	PN	PA	(MOLE-) MOLECULAR DYNAMICS INC.						XX PA
	XX	PI	Penn SG, Hanzel DK, Chen W, Rank DR;						XX PI
	XX	DR	WPI; 2001-488900/53.						XX DR
	XX	PT	Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow -						XX PT
	XX	PT	Example 4; SEQ ID NO: 27467; 650pp + Sequence Listing; English.						XX PS
	XX	XX	The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is a protein encoded by one of the probes of the invention.						XX CC
	XX	XX	Sequence 78 AA;						XX CC
	XX	XX	Query Match 48.9%; Score 135; DB 22; Length 78;						XX CC
	XX	XX	Best Local Similarity 53.6%; Pred. No. 3.2e-09;						XX CC
	XX	XX	Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;						XX CC
	XX	XX	Indels 4; Gaps 3;						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-2001; 2001WO-US00670.						XX CC
	XX	XX	30-OCT-2001 (first entry)						XX CC
	XX	XX	Peptide #1454, encoded by probe for measuring cervical gene expression.						XX CC
	XX	XX	Probe; human; microarray; gene expression; cervical epithelial cell; cervical cancer.						XX CC
	XX	XX	Homo sapiens.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	09-AUG-2001.						XX CC
	XX	XX	W0200157278-A2.						XX CC
	XX	XX	30-JAN-20						

XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-020456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632266.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-48897/53.
 XX Human Genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -
 XX PS Claim 27; SEQ ID No 27727; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENPs; see AAI13115/AI157546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders.
 XX PS Sequence 78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1495 encoded by probe for measuring placental gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;
 XX Query Match 48.9%; Score 135; DB 22; Length 78;
 CC Best Local Similarity 53.6%; Pred. No. 3.2e-09;
 CC Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
 CC Peptide #1432 encoded by probe for measuring breast gene expression.
 Qy 1 DNDDDDNDGGGGENN-GDGDNYATDD-MMNNMIMM-MVVRIMVVIMLM 52
 Db 12 DDDDDDDSDDDADNDGDDGDDGDDGDDGDDGDDGDDGDDGDDG 67
 XX OS Homo sapiens.
 XX PS Sequence .78 AA;

Predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human breast, particularly those diseases with polygenic aetiology. The diseases include: breast cancer, disorders of development, inflammatory diseases of the breast, fibrocystic changes, proliferative breast disease and non-carcinoma tumours.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained by electronic format directly from WIPO at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

RESULT 11
336623

ABG36823: ABG36823 Scutiger; Reptile; /8 AA.

19-AUG-2002 (First entry) Human peptide encoded by genome-derived single exon probe SEQ ID 26488.
Human; single exon probe; asthma; lung cancer; COPD; ILD; chronic obstructive pulmonary disease; interstitial lung disease; familial idiopathic pulmonary fibrosis; neurofibromatosis; tuberous sclerosis; Gaucher's disease; Niemann-Pick disease; Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis; pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome; pulmonary alveolar proteinosis; fibrocytic pulmonary dysplasia; primary ciliary dyskinesia; pulmonary hypertension; hereditary membranous disease; hereditary nephritis.

卷之三

卷之三

卷之三

200

卷之三

卷之三

卷之三

માનવ જીવનની પ્રક્રિયાઓ

C. N. SIMMERS AND R. H. INGERSOLL

תְּמִימָנָה וְעַמְלָה

-114183/15

of single exon nucleic acid probes used to addressable set

the application of the new technique -

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes also included in the array. The novel one

probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label, detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarray; having probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Kartagener Syndrome, fibrocytic pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

Note: The sequence data for this patent did not form part of the printed version of this patent, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published/pct/sequences](http://wipo.int/pub/published/pct/sequences).

SQ	Sequence	78	AA;
Query	Match	48.9%	Score 135; DB 23;
	Best Local Similarity	53.6%	Length 78;
	Matches	30;	Pred. No. 3 2e-09;
	Conservative	13;	Mismatches 9;
			Indels 4;
			Gaps
Qy	1	DNDGGDDNDGGGENNG-GDGNYATDD -MMNNMIMMM-MVVMVVMVYIMLM	52
Db	12	DDDDGGDDSDDDADNDGDDGDDGDDVTTMMVMMIIMAMMVMMVMMIVMM	67
RESULT 12			
aaacccggcc			

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

WO200175067-A2.
11-OCT-2001.

30-MAR-2001; 2001HQ-US08631.
31-MAR-2000; 2000US-0540217.
23-AUG-2000; 2000US-0649167.
(HYSE-) HYSEQ INC.

Sequence	402 AA;	Score 126; DB 22; Length 402;	Score 115.5; DB 22; Length 349;
Query Match	45.7%;	Best Local Similarity	41.8%;
Query Matches	27;	Pred. No.	4.1e-06;
Local Similarity	50.9%;	Conservative	16;
Matches	13;	Mismatches	10;
Conservative	13;	Indels	7;
Mismatches	9;	Gaps	Gaps
Indels	4;		3
Gaps	2;		
Qy	1 DNDD-GDDNNDGGGENGGDDNYATDD-MMMVMVIMVVIMLLM 52		
Db	107 DNGDVGDDNVDDDDGGDKDGDGGDTMVMMNMVLMVMMVVVVVV 165		

RESULT 14
 ABG19098
 ID ABG19098 Standard; Protein; 426 AA.
 XX
 AC ABG19098;
 XX
 DT 18-FEB-2002 (First entry)
 XX
 DE Novel human diagnostic protein #19089.
 XX
 KW Human; chromosome mapping; gene mapping; Gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; generic disorder;
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW diagnostic; generic disorder;
 XX

PCT PPT PPT PPT XX	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity	Claim 20; SEQ ID No 49457; 103pp; English.	<p>The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.</p> <p>Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://www.wipo.int/nrsub/pub/obligation.html</p>
--------------------------------	--	--	---

	Sequence	Match	Score	DB	Length	
xx	Sequence	426 AA;	41.3%;	114;	22;	426;
		Best Local Similarity	39.3%;	Pred. No.	7.9e-06;	
		Matches	22;	Mismatches	12;	Indels
		Conservative	18;			4
						Gaps
yy						1
bb						58

RESULT 15	
ABG48281	ABG48281 standard; Peptide; 105 AA.
X	
X	ABG48281;
X	
X	25-FEB-2003 (first entry)
X	
DB	Human liver peptide, SEQ ID NO 26939.
X	
X	Human; Liver; Atherosclerosis; hyperlipoproteinæmia; hyperlipidaemia;
X	hypercholesterolaemia; coronary heart disease;
X	hyperlipoproteinæmia; hyperlipidaemia;

Homo sapiens.
WO200157273-A2.

PD XX PP 09-AUG-2001.
30-JAN-2001; 2001WO-US00664

PR	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0608408.

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;

DR	WPI;	2001-488899/53.
XX	Human genome-derived	
PT	analysing gene expres-	
XX		
PS	Claim 27: SEQ ID No 2	
XX	The invention relates	
CC	measuring human gene	
CC	liver, comprising one	
CC	specification (or com-	
CC	stringency to a nucleic	
CC	liver. (I) may be used	
CC	expression in samples	
CC	identified may be inv-	
CC	hyperlipoproteinæmia	
CC	is associated with co-	
CC	human liver single ex-	
CC	Note: The sequence in	
CC	printed specification	
CC	from WIPO at ftp.wipo.	
XX		
SQ	Sequence 105 AA:	

Qy	2	NDDGDDND---
	:	
Db	5	DDYYDDDDGG
Qy	42	MVVVMVIMLM
	: : :	
Db	65	MVVVMVVVVVVVV

JOB TIME : 43 SECs

NUMBER OF SEQUENCES: 85
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 54570460 No. 5457046disk of No. 5457046 America, Inc.
 STREET: 405 Lexington Avenue, 62nd Floor
 CITY: New York
 STATE: United States of America
 COUNTRY: United States of America
 ZIP: 10174-6201
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/361.920
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/940,860
 FILING DATE: 28-OCT-1992
 APPLICATION NUMBER: DK 1158/90
 FILING DATE: 09-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/DK91/00124
 FILING DATE: 08-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 3435.204-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-867-0123
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 42 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 LENGTH: 42 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 LENGTH: 42 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 RESULT 3
 Query Match 25.0%; Score 69; DB 1; Length 42;
 Best Local Similarity 52.4%; Pred. No. 0.06; DB 1; Length 42;
 Matches 11; Conservative 6; Indels 4; Gaps 0;
 US-08-361-920-14

Query Match 25.0%; Score 69; DB 1; Length 42;
 Best Local Similarity 52.4%; Pred. No. 0.06; DB 1; Length 42;
 Matches 11; Conservative 6; Indels 4; Gaps 0;
 US-08-361-920-14

RESULT 4
 US-08-479-919-14

Sequence 14, Application US/08483432
 Patent No. 5763254

GENERAL INFORMATION:

APPLICANT: Weidike, Helle F.
 APPLICANT: Hagen, Frederick
 APPLICANT: Hjort, Carsten M.
 APPLICANT: Sven, Hasstrup
 TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
 NUMBER OF SEQUENCES: 85
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 56865930 No. 5686593disk of No. 5686593 America, Inc.
 STREET: 405 Lexington Avenue, 62nd Floor
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10174-6201

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/479,939
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/361,920
 FILING DATE: 22-DEC-1994
 APPLICATION NUMBER: US 07/940,860
 FILING DATE: 28-OCT-1992
 APPLICATION NUMBER: DK 1158/90
 FILING DATE: 09-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/DK91/00124
 FILING DATE: 08-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 3435.204-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-867-0123
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 42 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Humicola insolens
 STRAIN: DSM 1800
 US-08-479-919-14

Query Match 25.0%; Score 69; DB 1; Length 42;
 Best Local Similarity 52.4%; Pred. No. 0.06; DB 1; Length 42;
 Matches 11; Conservative 6; Indels 4; Gaps 0;
 Qy 2 NDDGDDNDGGGENNGGDN 22
 Db 3 SNNGGNNNGGNNNGGNN 23
 RESULT 4
 US-08-479-919-14
 Sequence 14, Application US/08483432
 Patent No. 5763254

GENERAL INFORMATION:

APPLICANT: Weidike, Helle F.
 APPLICANT: Hagen, Frederick
 APPLICANT: Hjort, Carsten M.
 APPLICANT: Sven, Hasstrup
 TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
 NUMBER OF SEQUENCES: 85
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 57632540 No. 5763254disk of No. 5763254 America, Inc.
 STREET: 405 Lexington Avenue, 62nd Floor
 CITY: New York
 STATE: New York
 COUNTRY: United States of America
 ZIP: 10174-6201

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/479,939
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/361,920
 FILING DATE: 22-DEC-1994
 APPLICATION NUMBER: US 07/940,860
 FILING DATE: 28-OCT-1992
 APPLICATION NUMBER: DK 1158/90
 FILING DATE: 09-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/DK91/00124
 FILING DATE: 08-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 3435.204-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEFAX: 212-867-0123
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 42 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Humicola insolens
 STRAIN: DSM 1800
 US-08-479-919-14

APPLICATION NUMBER: US/08/483,432
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/361,920
 FILING DATE: US 07/940,860
 APPLICATION NUMBER: US 07/940,860
 FILING DATE: 28-OCT-1992
 APPLICATION NUMBER: DK 1158/90
 FILING DATE: 09-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/DK91/00124
 FILING DATE: 08-MAY-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Lambiris, Elias J.
 REGISTRATION NUMBER: 33,728
 REFERENCE/DOCKET NUMBER: 3435.204-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 TELEX/FAX: 212-867-0298
 TELEFAX: 212-867-0298
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 42 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORGANISM: Humicola insolens
 STRAIN: DSM 1800
 US-08-483-432-14

Query Match 25.0%; Score 69; DB 1; Length 42;
 Best Local Similarity 52.4%; Pred. No. 0.06;
 Matches 11; Conservative 6; Mismatches 4; Indels 0;
 Gaps 0;
 Qy 2 NDDGDDNDGGGENNGGGDN 22
 Db 3 SNNGGNNNGGGNNNGGGNN 23

RESULT 5
 PC-US92-06840-2
 Sequence 2, Application PC/TUSS206840
 GENERAL INFORMATION:
 APPLICANT: Shi, Yang
 APPLICANT: Seto, Edward
 APPLICANT: Shenk, Thomas
 TITLE OF INVENTION: YY1 TRANSCRIPTION FACTOR AND METHODS OF
 NUMBER OF SEQUENCES: 10
 CURRENT APPLICATION DATA:
 ADDRESSSEE: Ostvollenk, Faber, Gerb & Soffen
 STREET: 1180 Avenue of the Americas - 7th Floor
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-8403

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/06840
 FILING DATE: 19920814
 CLASSIFICATION: AU 1805
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/746,485
 FILING DATE: 16-AUG-1991
 ATTORNEY/AGENT INFORMATION:

NAME: Dennis, Manette
 REGISTRATION NUMBER: 30,623
 REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 382-0700
 TELEFAX: (212) 382-0888
 TELEX: 23625
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 414 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US92-06840-2
 Query Match 25.0%; Score 69; DB 1; Length 414;
 Best Local Similarity 45.8%; Pred. No. 0.88;
 Matches 11; Conservative 5; Mismatches 8; Indels 0;
 Gaps 0;
 Qy 1 DNDDGDDNDGGGENNGGGDNYA 24
 Db 45 EEEDDDDGGGDGGGGGGHGA 68
 RESULT 6
 US-08-681-129-2
 Sequence 2, Application US/08681129
 GENERAL INFORMATION:
 APPLICANT: Mettenleiter, Thomas Christoph
 TITLE OF INVENTION: Pseudorabies virus vaccine
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Organon Teknika Corporation
 STREET: 1130 Piccard Drive
 CITY: Rockville
 STATE: Maryland
 COUNTRY: U.S.A.
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/681,129
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/244,446
 FILING DATE: 02-JUN-1994
 CLASSIFICATION: 424
 APPLICATION NUMBER: EP 92.203.079.6
 FILING DATE: 06-OCT-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Gorwley, Mary E.
 REGISTRATION NUMBER: 34,409
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 238-5200
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 525 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Pseudorabies virus
 STRAIN: Kaplan
 US-08-681-129-2
 Query Match 24.1%; Score 66.5; DB 1;
 Best Local Similarity 28.3%; Pred. No. 2.3;

RESULT 7

US-08-375-300-4
Sequence 4, Application US/08315300
Patent No. 5679366

GENERAL INFORMATION:
APPLICANT: Feng, He
ATTORNEY: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6

ADDRESSSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-375-300-4

Query Match 23.9%; Score 66; DB 1; Length 764;
Best Local Similarity 48.3%; Matches 5; Mismatches 8; Indels 2; Gaps 1;

RESULT 9

PCT-US95-16930-4
Sequence 4, Application PC/TUS9516930

GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY
TITLE OF INVENTION: FUNCTION
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16930

FILING DATE: 27-DEC-1995
PRIORITY NUMBER: US 08/375,300

APPLICANT: Feng, He
ATTORNEY: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:
ADDRESSSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

RESULT 7

US-08-375-300-4
Sequence 4, Application US/08315300
Patent No. 5679366

GENERAL INFORMATION:
APPLICANT: Feng, He
ATTORNEY: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 6

ADDRESSSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 764 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide

US-09-177-431-4

Query Match 23.9%; Score 66; DB 3; Length 764;
Best Local Similarity 48.3%; Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy 1 DNDDGDDNDG - GGENNGGDNYATDD 27
Db 574 DDDDDDDDGEGDDDEDDDEDD 602

RESULT 8

US-09-177-431-4
Sequence 4, Application US/09177431
Patent No. 6071700

GENERAL INFORMATION:
APPLICANT: Feng, He
ATTORNEY: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:
ADDRESSSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

REGISTRATION NUMBER: 32,983
REFERENCE DOCKET NUMBER: 04020/046W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906

TELEX: 200154 INFORMATION FOR SEQ ID NO: 4
 SEQUENCE CHARACTERISTICS:
 LENGTH: 764 amino acids
 TYPE: amino acid
 STRANEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Protein PCT-US95-16930-4

Query Match 23.9%; Score 66; DB 5; Length 764;
 Best Local Similarity 48.3%; Pred. No. 4.1;
 Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Db 574 DNDDDDNDDG-GGENNGGDDNYATDD 27
 Db 574 DDDDDDDDGEGDDEDDDEDDDD 602

RESULT 10 US-08-375-300-2
 Sequence 2, Application US/08375300
 Patent No. 5679566
 GENERAL INFORMATION:
 APPLICANT: Feng, He
 ATTORNEY/AGENT: Jacobson, Allan S.
 TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street Suite 3100
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 LENGTH: 1089
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/375,300
 FILING DATE: 20-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Fasse, J. P.
 REGISTRATION NUMBER: 32,983
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEX: (617) 542-8906
 TELEX: 200154 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1089 amino acids
 STRANEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein US-08-375-300-2

Query Match 23.9%; Score 66; DB 1; Length 1089;
 Best Local Similarity 49.3%; Pred. No. 6.3;
 Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Db 899 DNDDDDNDDG-GGENNGGDDNYATDD 27
 Db 899 DDDDDDDDGEGDDEDDDEDDDD 927

RESULT 11 US-09-177-411-2
 Sequence 2, Application US/09177431

Query Match 23.9%; Score 66; DB 3; Length 1089;
 Best Local Similarity 48.3%; Pred. No. 6.3;
 Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Db 899 DNDDDDNDDG-GGENNGGDDNYATDD 27
 Db 899 DDDDDDDDGEGDDEDDDEDDDD 927

RESULT 12 PCT-US95-16930-2
 Sequence 2, Application PC/TUS9516930
 GENERAL INFORMATION:
 APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
 TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street Suite 3100
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 LENGTH: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09177431
 FILING DATE: 20-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Fasse, J. P.
 REGISTRATION NUMBER: 32,983
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEX: (617) 542-8906
 TELEX: 200154 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1089 amino acids
 STRANEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein US-09-177-411-2

RESULT 13 Sequence 2, Application PC/TUS9516930-2

APPLICATION NUMBER: PCT/US95/16930
 FILING DATE: 27-DEC-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/375,300
 FILING DATE: 20-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Fasse, J. Peter
 REGISTRATION NUMBER: 32,983
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)542-5070
 TELEX: (617)542-8906
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1089 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-16930-2

Query Match Score 66; DB 5; Length 1089;
 Best Local Similarity 48.3%;
 Matches 14; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

Qy 1 DNDDGDDDDDG - GGEENNDDGNYATD 27
 Db 899 DDDDDDDDDGEGEDDDDDDEDDDD 927

RESULT 13
 US-08-560-398-8
 Sequence 8, Application US/08560398
 Patent No. 5907082
 GENERAL INFORMATION:
 APPLICANT: O'Neill, Sharman
 APPLICANT: Nadeau, Jeanette
 TITLE OF INVENTION: Ovule-Specific Gene Expression
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESS: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 CITY: San Francisco
 STATE: California
 ZIP: 94111-3034
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560,398
 FILING DATE: 17-NOV-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Bastian, Kevin L.
 REGISTRATION NUMBER: 34,774
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 190 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-398-8

Query Match Score 65.5; DB 2; Length 190;
 Best Local Similarity 46.4%;
 Pred. No. 0.93;

Matches 13; Conservative 2; Mismatches 6; Indels 7; Gaps 1;

Qy 1 DNDDGDD-----DNDDGGGENGGGD 21
 Db 109 DNGEGDSDQPIGSGNNDGNGNDGNSD 136

RESULT 14
 US-08-258-188-2
 Sequence 2, Application US/08258188
 Patent No. 5475098
 GENERAL INFORMATION:
 APPLICANT: Hall, Robert H.
 APPLICANT: XU, Jian Guo
 TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E. coli O157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND SPECIFIC DETECTION OF O157:H7 AND OTHER ENTEROHEMORRHAGIC COLI
 TITLE OF INVENTION: A New and Distinctive DNA Sequence of E. coli O157:H7 and its Use for the Rapid, Sensitive and Specific Detection of O157:H7 and Other Enterohemorrhagic Coli
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Khourie and Crew
 STREET: Stewart Street Tower, One Market Plaza
 CITY: San Francisco
 STATE: California
 COUNTRY: US
 ZIP: 94105-1493
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/258,188
 FILING DATE: 14-JUN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Garrett-Wackowski, Eugenia
 REGISTRATION NUMBER: 37,330
 REFERENCE/DOCKET NUMBER: 15280206, DHHS/SE135540
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 543-9600
 TELEXFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 758 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-258-188-2

Query Match Score 65; DB 1; Length 758;
 Best Local Similarity 48.6%;
 Matches 17; Conservative 2; Mismatches 4; Indels 12; Gaps 3;

Qy 3 DDGDD-----NDD---GGEN--NCGDGNYAT 25
 Db 505 DDGDDYTSGGGDDQLFGSGNDKLSCGDGNNYLT 539

RESULT 15
 US-08-526-813-2
 Sequence 2, Application US/08526813
 Patent No. 5756293
 GENERAL INFORMATION:
 APPLICANT: Hall, Robert H.
 APPLICANT: XU, Jian Guo
 TITLE OF INVENTION: A New and Distinctive DNA Sequence of E. coli O157:H7 and its Use for the Rapid, Sensitive and Specific Detection of O157:H7 and Other Enterohemorrhagic Coli
 Patent No. 5756293
 TITLE OF INVENTION: colI O157:H7 and its Use for the Rapid, Sensitive and Specific Detection of O157:H7 and Other Enterohemorrhagic Coli
 TITLE OF INVENTION: A New and Distinctive DNA Sequence of E. coli O157:H7 and its Use for the Rapid, Sensitive and Specific Detection of O157:H7 and Other Enterohemorrhagic Coli
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
COMPUTER TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/526,813
FILING DATE: 11-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,188
FILING DATE: 14-JUN-1994

ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 15280-206-1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 758 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-526-813-2

Query Match Similarity 23.6%; Score 65; DB 1; Length 758;
Best Local Matches 48.6%; Pred. No. 5,4;
Matches 17; Conservative 2; Mismatches 4;
Indels 12; Gaps 3;

Qy 3 DDGND-----NND---GGEN---NGGDGDNYAT 25
Db 505 DDCDDYISGGDQLFEGSGNDLSSGGGNYYLT 539

Search completed: December 3, 2003, 16:20:39
Job time : 22 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003 16:19:35 ; Search time 30 Seconds (without alignments)
328.571 Million cell updates/sec

Title: US-09-991-003B-8

Perfect score: 276

Scoring table: BL0SUM62

Sequence: 1 DNDDDDNDGGGENGGDG.....MMVVVRIMVMMVIMLLMT 53

Searched: 684280 seqs, 185583659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 45 summaries .

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

3: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

4: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

6: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep:*

8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpa/US10 NEW_PUB.com:*

17: /cgn2_6/ptodata/1/pubpa/US60 NEW_PUB.com:*

18: /cgn2_-/ptodata/1/pubpa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query ID	Match ID	Length	DB ID	Description
1	276	100.0	53	10	US-09-991-003B-8	Sequence 8, Appli
2	135	48.9	78	9	US-09-864-761-34745	Sequence 34745, A
3	121	43.8	315	12	US-10-017-161-1914	Sequence 1914, Ap
4	109	39.5	105	9	US-09-864-761-34191	Sequence 34191, A
5	108.5	39.3	276	12	US-10-017-161-2048	Sequence 2048, Ap
6	100	36.2	410	12	US-10-017-161-1820	Sequence 1820, Ap
7	96.5	35.0	329	12	US-10-017-161-1782	Sequence 1782, Ap
8	93.5	33.9	53	9	US-09-864-761-36649	Sequence 36649, A
9	91	33.0	138	9	US-10-029-386-30443	Sequence 30443, A
10	89.5	32.4	52	12	US-10-029-386-30443	Sequence 30443, A
11	80.5	29.2	195	9	US-09-864-761-33537	Sequence 33537, A
12	80	29.0	69	9	US-09-864-761-42285	Sequence 42285, A
13	80	29.0	255	12	US-10-017-161-2180	Sequence 2180, Ap
14	77.5	28.1	59	12	US-10-029-386-7424	Sequence 27424, A
15	77	27.9	12	US-10-029-386-314083	Sequence 34083, A	
16	77	27.9	125	9	US-09-864-761-37222	Sequence 37222, A
17	77	27.9	173	9	US-09-864-761-16371	Sequence 36371, A
18	77	27.9	173	12	US-10-029-386-32638	Sequence 33238, A
19	77	27.5	109	9	US-09-864-761-13541	Sequence 33541, A
20	76	27.5	75	9	US-09-864-761-47141	Sequence 47141, A
21	75	27.2	75	6	US-10-029-386-31671	Sequence 34171, A
22	74	26.8	60	12	US-09-864-761-44209	Sequence 44209, A
23	74	26.1	102	9	US-09-864-761-44209	Sequence 44209, A
24	71.5	25.9	450	12	US-10-017-161-2050	Sequence 2050, Ap
25	71	25.7	103	9	US-10-029-386-31210	Sequence 44210, A
26	71	25.5	235	12	US-10-029-386-32032	Sequence 32032, A
27	70.5	25.4	48	9	US-09-864-761-36445	Sequence 36445, A
28	70	25.4	24	12	US-10-029-386-3037	Sequence 3037, A
29	70	25.4	357	9	US-09-864-761-35807	Sequence 35807, A
30	69.5	25.2	66	9	US-09-864-761-33961	Sequence 33961, A
31	69	25.0	178	12	US-10-029-386-33408	Sequence 33408, A
32	68.5	24.8	31	12	US-10-029-386-33523	Sequence 33523, A
33	68.5	24.8	207	10	US-09-965-602-34	Sequence 34, App1
34	68	24.6	50	9	US-09-864-761-44258	Sequence 44258, A
35	67.5	24.5	35	12	US-10-029-386-33525	Sequence 33525, A
36	67.5	24.5	44	12	US-10-029-386-31672	Sequence 31672, A
37	67	24.3	32	9	US-09-864-761-15019	Sequence 35019, A
38	67	24.3	722	12	US-10-161-051-193	Sequence 151, App
39	66	23.9	41	9	US-09-864-761-34567	Sequence 34567, A
40	65.5	23.7	371	10	US-09-738-626-5358	Sequence 5358, Ap
41	65	23.6	28	9	US-09-864-761-35917	Sequence 35917, A
42	65	23.6	63	9	US-09-864-761-34909	Sequence 34409, A
43	65	23.6	762	10	US-09-738-626-4825	Sequence 4825, Ap
44	64.5	23.4	936	12	US-10-032-585-7688	Sequence 7688, Ap
45	64.5	23.4	1230	12	US-10-205-219-44	Sequence 44, App1

ALIGNMENTS

RESULT 1
US-09-991-003B-8
; Sequence 8, Application US-09991003B
; SEQ ID NO. US200201712541
; GENERAL INFORMATION:
; APPLICANT: KAMB, Cary Alexander
; APPLICANT: PORITZ, Mark Aaron
; APPLICANT: TENG, David Heng-Fai
; TITLE OF INVENTION: Human Rhinovirus Assays, and Compositions Therefrom
; FILE REFERENCE: 29145/36971A
; CURRENT APPLICATION NUMBER: US-09-991-003B
; CURRENT FILING DATE: 2002-11-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: perturbagen_sequence
; LOCATION: (1)..(53)
; OTHER INFORMATION: cwn85 perturbagen
; US-09-991-003B-8

RESULT 2
US-09-864-761-34745
; Sequence 4, Application US-09864761
; SEQ ID NO. US-00048763A1
; GENERAL INFORMATION:

```

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00660
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Amomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 34745
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL080246.13
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN FET74, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; US-09-864-761-34745

; Query Match 48,91; Score 135; DB 9; Length 78;
; Best Local Similarity 53.64%; Pred. No. 4.e-08;
; Matches 30; Conservative 13; Mismatches 9; Indels 4; Gaps 3;
; 1 DNDGDDNDGGCENNG-GGDGYATDD--MMVVMIMM-MVVVRIMMVVMVNMLL 52
; 12 DDGDDGDDSDDDDDNDGCGDGDGDDDVMTMVMAMVMMVMMVMM 67

```

PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00660
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09//774,203
 PRIOR FILING DATE: 2001-01-29
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 34191
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AL018584.17
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 11
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 11
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8.8
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.6
 US-09-864-761-34191

Query Match Score 109; DB 9; Length 105;
 Best Local Similarity 39.5%; Pred. No. 4.5e-05;
 Matches 25; Conservative 16; Indels 20; Gaps 3;

Qy 2 NDDGDDND-----DGDRYATDDMMVMMVMMVV-----RI 41
 Db 5 DDDYDDTDDGGGSWTCGSDSSGGSDYRDNNNCGGKRMIIIMMIVVVVLLVMMV 64

Qy 42 MVVMVITIMLUM 52
 Db 65 MVVMVVMVMMV 75

RESULT 6
 US-10-017-161-1820
 ; Sequence 1820, Application US/10017161
 ; Publication No. US2003014368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUWA, MAKIKO
 ; APPLICANT: ASAI, KIYOSHI
 ; APPLICANT: AKIYAMA, YUTAKA
 ; APPLICANT: ABURATANI, HIROYUKI
 ; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 ; FILE REFERENCE: 08435/0152
 ; CURRENT APPLICATION NUMBER: US/10/017,161
 ; CURRENT FILING DATE: 2002-12-18
 ; PRIOR APPLICATION NUMBER: JP 2001/246789
 ; NUMBER OF SEQ ID NOS: 2430
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1820
 ; LENGTH: 410
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-017-161-1820

Query Match Score 100; DB 12; Length 410;
 Best Local Similarity 36.2%; Pred. No. 0.0018;
 Matches 25; Conservative 13; Mismatches 12; Indels 40; Gaps 2;

Qy 3 DDGDDNDGGGENNG-----GDGDNYATDDMMVMM 32
 Db 115 DEGGGGDCCGNDGDDGYCSGSVDHGGSDGGHDVGGVGVGIGVGDGGCNFHGMIV 174

Qy 33 IM-----MVVMVITIMLUM 52
 Db 175 VVVVTTMTVVEGMVVMVIVVVVTVLVV 204

RESULT 7
 US-10-017-161-1782
 ; Sequence 1782, Application US/10017161
 ; Publication No. US2003014368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUWA, MAKIKO
 ; APPLICANT: ASAI, KIYOSHI
 ; APPLICANT: AKIYAMA, YUTAKA
 ; APPLICANT: ABURATANI, HIROYUKI
 ; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 ; FILE REFERENCE: 08435/0152
 ; CURRENT APPLICATION NUMBER: US/10/017,161
 ; CURRENT FILING DATE: 2002-12-18
 ; PRIOR APPLICATION NUMBER: JP 2001/246789
 ; NUMBER OF SEQ ID NOS: 2430
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1782
 ; LENGTH: 329
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-017-161-1782

Query Match Score 96.5%; DB 12; Length 329;
 Best Local Similarity 35.0%; Pred. No. 0.0036;
 Matches 21; Conservative 14; Mismatches 9; Indels 11; Gaps 2;

Query Match Score 108.5; DB 12; Length 276;
 Best Local Similarity 39.3%; Pred. No. 0.00014; Mismatches 18; Indels 11; Gaps 1;

Qy 2 NDDGDDNDGGENNGDDNYATDDMMVVMM-----MMVVRIMVVMVIM 49
 Db 73 DDDGSSDGGGGD----DGGVVVVLLVVVVMAVMGVVVVVVVVVVVVVVV 123

RESULT 8
 US-09-864-761-34192
 ; sequence 34192, Application US/09864761
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Acomica-X-1
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-10
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR FILING DATE: 2001-01-30
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR FILING NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Ammax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 34192
 LENGTH: 53
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:

OTHER INFORMATION: MAP TO AL07854.17
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 11
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8.8
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.6

Qy 2 NDDGDDN-DDGGGENN----GGGGNATDDMMVVMMVV 38
 Db 10 DDDYDNVDDGGDGSYSCGGDGSGGEVMMIETMMVV 53

RESULT 9
 US-09-864-761-36649
 ; Sequence 36649, Application US/09864761
 ; Patent No. US20020048764A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Acomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 09/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR FILING DATE: 2000-06-30
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Ammax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 36649
 LENGTH: 138
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:

OTHER INFORMATION: MAP TO AL035419.9
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
 US-09-864-761-36649

Query Match Score 91; DB 9; Length 138;
 Best Local Similarity 59.3%; Pred. No. 0.0059; Indels 0; Gaps 0;
 Matches 16; Conservative 3; Mismatches 8;

Qy 1 DNDGDDNDGGENNGGDNYATDD 27
 Db 108 DGDGDDNNDDGDDATAVATDD 134

RESULT 10
 US-10-029-386-30443
 Sequence 30443, Application US/10029386
 Publication No. US2010194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
 FILE REFERENCE: AEMOTICA-X-2
 CURRENT APPLICATION NUMBER: US/10/029,386
 CURRENT FILING DATE: 2010-12-20
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
 SEQ ID NO 30443
 LENGTH: 52
 TYPE: PRT
 ORGANISM: Homo sapiens

FEATURE:
 OTHER INFORMATION: MAP TO AC015478.3
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
 US-10-029-386-30443

Query Match Score 89.5; DB 12; Length 52;
 Best Local Similarity 69.6%; Pred. No. 0.0031; Indels 1; Gaps 1;
 Matches 16; Conservative 4; Mismatches 2;

Qy 1 DNDGDDNDGGENNG-GDGN 22
 Db 108 DDDDDGDSDDGGDNNGSSGDD 50

RESULT 11
 US-09-864-761-33537
 Sequence 33537, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
 FILE REFERENCE: Aemotica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-03-03
 PRIOR APPLICATION NUMBER: GB 24263,6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-05-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 NUMBER OF SEQ ID NOS: 49117
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
 SEQ ID NO 33537
 LENGTH: 195

TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.14
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.8
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.3
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.2
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.8
 OTHER INFORMATION: EXPRESSED IN BT74, SIGNAL = 7
 US-09-864-761-33537

Query Match Score 29.2%; DB 9; Length 195;
 Best Local Similarity 34.7%; Prd. No. 0.12; Mismatches 10; Indels 13; Gaps 1;

Qy 4 DGDDNDGGENNGGDGNYATDDMMMMVVVIMVLM 52
 Db 2 DGCHGDDGAAGN-----VVVVVVVVVVVVVVVVVV 37

RESULT 12
 US-09-864-761-42285
 Sequence 42285, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
 FILE REFERENCE: Aemotica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
 FILE REFERENCE: Aemotica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26

FILE REFERENCE: Aemotica-X-1

PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: AnnoMax Sequence Listing Engine vers. 1.1
 SEQ ID NO 35629
 LENGTH: 42
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC009492.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
 US-09-864-761-35629

Query Match 27.9%; Score 77; DB 9; Length 42;
 Best Local Similarity 50.0%; Pred. No. 0.06;
 Matches 13; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 DNDDGDDNDGGGENNGGDNATD 26
 Db 17 DBSGDDDDDDNDDGSDDDVGD 42

Result No.	Score	Query	Match	Length	DB ID	Description
1	79	BLOSUM62	SPBP RAT	279	1	P08723 rattus norvegicus (Rat) SPBP protein
2	74	Gapop 10.0 , Gapext 0.5	CAQ2_MOUSE	415	1	P09161 mus musculus CAQ2 protein
3	73.5	Searched: 127863 seqs, 47026705 residues	NNP1_MOUSE	494	1	P56183 mus musculus NNP1 protein
4	73	Total number of hits satisfying chosen parameters:	YAB9_YEAST	1131	1	P31380 saccharomyces cerevisiae YAB9 protein
5	70	Minimum DB seq length: 0	VG11_HSV1	318	1	P00120 iceluridoh VG11 protein
6	70	Maximum DB seq length: 2000000000	IB68_PRVKA	364	1	P24827 pseudorabies virus IB68 protein
7	70	Post-processing: Minimum Match 0%	YGG6_YEAST	657	1	P53165 saccharomyces cerevisiae YGG6 protein
8	69	Listing first 45 summaries	PRP3_RAT	206	1	P04474 rattus norvegicus PRP3 protein
9	69	Database : SwissProt_41; *	TYY1_HUMAN	414	1	P25430 homo sapiens TYY1 protein
10	68.5	68.5	WJY2_YEAST	207	1	P47087 saccharomyces cerevisiae WJY2 protein
11	68	68.5	KEK2_CANAL	924	1	P013359 candida albicans KEK2 protein
12	67.5	67.5	GP10_DICDI	544	1	Q06885 dictyospermi GP10 protein
13	67	66.5	CAQ2_RABBIT	409	1	P31235 orctolagus cuniculus CAQ2 protein
14	66.5	66.5	TR16_POPTR	340	1	P16579 populus tremuloides TR16 protein
15	66.5	66.5	CAQ2_RAT	413	1	P51868 rattus norvegicus CAQ2 protein
16	66	66	YNL5_YEAST	644	1	P53935 saccharomyces cerevisiae YNL5 protein
17	66	66	NMD2_YEAST	1089	1	P38198 saccharomyces cerevisiae NMD2 protein
18	66	66	YB010_YEAST	1094	1	P38194 vaccinia virus YB010 protein
19	64	64	VF05_VACC	321	1	P21014 vaccinia virus VF05 protein
20	64	64	VFO5_VACC	322	1	P24358 vaccinia virus VFO5 protein
21	64	64	TYY1_MOUSE	414	1	P05059 variola virus TYY1 protein
22	63	63	VFO5_VARV	322	1	P33865 variola virus VFO5 protein
23	63	63	CSP_FLAWS	442	1	P08307 plasmodium CSP protein
24	63	63	COAA_BP1F1	460	1	Q08027 bacteriophaga COAA_BP1F1 protein
25	63	63	KEX11_YEAST	729	1	P09520 saccharomyces cerevisiae KEX11 protein
26	63	63	VHT11_YEAST	840	1	P388155 saccharomyces cerevisiae VHT11 protein
27	63	63	YNF4_YEAST	1165	1	P53350 saccharomyces cerevisiae YNF4 protein
28	62.5	62.5	TBG_CANAL	502	1	Q93007 candida albicans TBG_CANAL protein
29	62	62	GRPA_MEDICA	159	1	Q91334 medicago sativa GRPA_MEDICA protein
30	62	62	RMP_HUMAN	508	1	Q94163 homo sapiens RMP_HUMAN protein
31	62	62	EBNI_EBV	641	1	P03211 Epstein-Barr virus EBNI protein
32	61.5	61.5	COAA_BP1F1	437	1	P541329 bacteriophage COAA_BP1F1 protein
33	61.5	61.5	MM15_MOUSE	657	1	P541322 mus musculus MM15 protein

ALIGNMENTS						
RESULT 1						
SPBP_RAT						
ID	SPBP_RAT	STANDARD;	PRT;	279 AA.		
AC	P08723;					
DT	01-AUG-1988 (Rel. 08, Created)					
DT	01-AUG-1988 (Rel. 08, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DE	Prostatic spermine-binding protein precursor (SPBP).					
GN	SPBP.					
OS	Rattus norvegicus (Rat).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
[1]						
RN	SEQUENCE FROM N.A.					
RP	RPIDIINE=87137538; PubMed=3810623;					
RX	Anderegg R.J., Carr S.A., Huang I.-Y., Hiipakka R.A., Chang C., Liao S.;					
RA	"Correction of the cDNA-derived protein sequence of prostatic spermine binding protein: pivotal role of tandem mass spectrometry in sequence analysis".					
RT	Biochemistry 27:4214-4221(1998).					
RT	-1- FUNCTION: SPERMINE-BINDING PROTEIN IS AN ANDROGEN REGULATED VENTRAL PROSTATE GLYCOPROTEIN THAT BINDS VARIOUS POLYAMINES.					
RL	-1- TISSUE SPECIFICITY: Prostate.					
CC	-1- SIMILARITY: TO MOUSE SPBP.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC	EMBL; J02675; AAA44113.1; ALT SEQ					
CC	InterPro; IPR01229; Jacalin_Lectin.					
CC	Pfam; PF01419; Jacalin_1.					
DR	DR IntercPro; IPR01229; Jacalin_Lectin.					
KW	KW Glycoprotein; Signal; Pyrrolidone carboxylic acid.					
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279	PROSTATIC SPERMINE-BINDING PROTEIN PYRROLIDONE CARBOXYLIC ACID.		
FT	FT MOD RES	18	18	N-LINKED (GLCNAC-).		
FT	FT CAREHYD	62	62	ASP/GLU-RICH (ACIDIC).		
FT	FT DOMAIN	154	279	ASP/GLU-RICH (ACIDIC).		
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT	FT CHAIN	18	279			
FT	FT MOD RES	18	18			
FT	FT CAREHYD	62	62			
FT	FT DOMAIN	154	279			
FT	FT SEQUENCE	279 AA;	31000 MW;	3BBB01A02A517A65 CRC64;		
FT	FT SIGNAL	1	17			
FT						

Qy	1 DNDGDD--NDDGGENNNGGDGNYATDD 27	Oy	1 DNDGDDDDNDDGGGENNGGDGNYATDD 27
Db	239 DNDDEDDDDDDGGSSDDGGDDDDDD 267	Db	383 DGDDNNDDDDDDNNSDEDNEEDSDD 409
RESULT 2			
CAQ2_MOUSE	CAQ2 MOUSE STANDARD; PRT: 415 AA.	CAQ2_MOUSE	CAQ2 MOUSE STANDARD; PRT: 494 AA.
AC	009161; 088505; (Rel. 35, Created)	ID	NNP1_MOUSE
DT	01-NOV-1997 (Rel. 38, Last sequence update)	AC	PS6183; Q35712; Q9JKE7; Q9JKU2;
DT	15-JUL-1999 (Rel. 38, Last annotation update)	DT	01-NOV-1997 (Rel. 35, Created)
DT	28-FEB-2003 (Rel. 41, Last annotation update)	DT	28-FEB-2003 (Rel. 41, Last sequence update)
DE	Calsequestrin, cardiac muscle isoform precursor (Calsequestrin 2).	DE	NNP-1 protein (Novel nuclear protein 1) (Nop52).
DN	Mus musculus (Mouse).	GN	NNP1.
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	OS	Mus musculus (Mouse).
OC	[1]	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;	OX	NCBI_TaxID=10090;
RN		RN	SEQUENCE FROM N.A. (ISOFORM 2), AND DEVELOPMENTAL STAGE.
RP	SEQUENCE FROM N.A. (ISOFORM 2), AND DEVELOPMENTAL STAGE.	RP	SEQUENCE FROM N.A. (ISOFORM 2), AND DEVELOPMENTAL STAGE.
RX	MEDLINE:99014233; PubMed=9751116;	RC	TISSUE-Pituitary;
RA	Park K.W., Goo J.H., Chung H.-S., Kim H., Kim D.-H., Park W.-J.; RT	RA	Medline:97336061; PubMed=9192856;
RT	"Cloning of the genes encoding mouse cardiac and skeletal calsequestrin: expression pattern during embryogenesis.";	RA	Jansen E., Meulemans S.M.P., Orlans I.C.R., Van de Ven W.J.M.; RT
RT	Gene 21:25-30 (1998).	RT	"The NNP-1 gene (D21S2056E), which encodes a novel nuclear protein maps in close proximity to the cystatin B gene within the EPM1 and APECD critical region on 21q22.3.";
RN	[2]	RT	Genomics 42:336-341 (1997).
RP	SEQUENCE FROM N.A.	RL	APECD critical region on 21q22.3.";
RC	SEQUENCE FROM N.A. (ISOFORM 1).	RN	[2]
RA	Sato Y., Ferguson D.G., Sako H., Dorn G.W. II, Kadambi V.J., RT	RC	SEQUENCE FROM N.A. (ISOFORM 1).
RA	Yatani A., Hoit B.D., Walsh R.A., Kranias E.G.; RT	RA	Jansen E.;
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.	RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC	-!- CALCIUM-BINDING PROTEIN IS A HIGH-CAPACITY, MODERATE AFFINITY, CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INDIRECT CALCIUM STORE IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40 TO 50 MOLES OF CALCIUM (BY SIMILARITY).	CC	-!- CALCIUM-BINDING PROTEIN IS A HIGH-CAPACITY, MODERATE AFFINITY, CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INDIRECT CALCIUM STORE IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40 TO 50 MOLES OF CALCIUM (BY SIMILARITY).
CC	-!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF CARDIAC AND SLOW SKELETAL MUSCLE CELLS.	CC	-!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF CARDIAC AND SLOW SKELETAL MUSCLE CELLS.
CC	-!- TISSUE SPECIFICITY: HEART.	CC	-!- TISSUE SPECIFICITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
CC	-!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.	CC	-!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL: U91483; AAC2512.1; -.	CC	Event=Alternative splicing; Named isoforms=3;
DR	EMBL: AF068244; AAC69472.1; -	CC	Name=1; IsoId=P56183-1; Sequence=Displayed;
DR	HSSP; P0721; 1A8V.	CC	Name=2; IsoId=P56183-2; Sequence=vSP_004335, VSP_004336;
DR	MGB; MG1:130469; Casq2.	CC	Name=3; IsoId=P56183-3; Sequence=vSP_004337;
DR	InterProc: IP001193; Calsequestrin.	CC	Event=Alternative splicing; Named isoforms=3;
DR	PRINTS: PR0126; Calsequestrin; -.	CC	Event=Alternative splicing; Named isoforms=3;
DR	PROSITE: PS00863; CALSEQUESTRIN_1; 1.	CC	Event=Alternative splicing; Named isoforms=3;
DR	PROSITE: PS00864; CALSEQUESTRIN_2; 1.	CC	Event=Alternative splicing; Named isoforms=3;
KW	Muscle; Glycoprotein; Calcium-binding. Signal.	CC	Event=Alternative splicing; Named isoforms=3;
FT	SIGNAL 1 19 POTENTIAL.	CC	Event=Alternative splicing; Named isoforms=3;
FT	CHAIN 20 415 CALSEQUESTRIN, CARDIAC MUSCLE ISOFORM.	CC	Event=Alternative splicing; Named isoforms=3;
FT	DOMAIN 373 415 ASP GLU-RICH (ACIDIC).	DR	EMBL: U79773; AAC53286.1; -.
FT	CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).	DR	EMBL: U79774; AAC63261.1; -.
FT	CONFLICT 1 2 Q -> E (IN REF. 1).	DR	EMBL: AF294729; AAC76216.2; -.
FT	CONFLICT 20 20 H -> D (IN REF. 2).	DR	EMBL: AF312394; AAC30393.1; -.
FT	CONFLICT 278 278 MGI:1203500; Nnp1.	DR	MGB; MGI:1203500; Nnp1.
SQ	SEQUENCE 415 AA; 48197 MW; A397ED046F8BD7E CRC4;	KW	Nuclear protein; Alternative splicing.
FT	Score 74 DB 1; Length 415;	FT	ASPG-RICH.
FT	Pred. No. 0.27; Indels 0; Gaps 0;	FT	ASGCDGFSASGDD -> GSDEDDDDDSDEA (in 1b)form 2).
FT	4; Mismatches 13; Conservative Matches 13;	FT	/FTId=vSP_004335.

PT	VARSPLIC	259	309	Missing (in isoform 2). /FTId=VSP_004336.	DR	InterPro; IPR001410; DEAD.
PT	VARSPLIC	450	484	Missing (in isoform 3). /FTId=VSP_004337.	DR	InterPro; IPR001650; Helicase_C.
PT	CONFLICT	50	50	D -> C (IN REF. 3; AAG302931).	DR	InterPro; IPR000330; SNF2_N.
SQ	SEQUENCE	494 AA;	54776 MW;	9886F1BB7AE3E99D CRC64;	DR	Pfam; PF00271; helicase_C; 1.
Query Match		26 6%; Best Local Similarity 53.6%; Matches 15; Conservative 5; Mismatches 7; Indels 1; Gaps 1;	Score 73.5; DB 1; Length 494; Pred. No. 0.36; PRT; 1131 AA.	DR	SMART; SM00467; DEXDC; 1.	
Qy	1	DNDGDDDDGGCG-NNGGDNYATDD 27	NP-BIND ATP (POTENTIAL).	DR	SMART; SM00490; HELICC; 1.	
Db	255	DGDGEASDDDGAEASDGSDGVADSD 282	FT SITE 703 DEGH BOX.	DR	Hydrothetical protein; Nuclear protein; DNA-binding; Helicase; KW ATP-binding.	
Query Match		26.4%; Best Local Similarity 52.0%; Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;	Score 73; DB 1; Length 1131; Pred. No. 0.95; PRT; 1131 AA.	DR	Nuclear protein; DNA-binding; Helicase; KW ATP-binding.	
RESULT 4	YAB9_YEAST	STANDARD:	DB 407 DDNDNDGNNNNNNNNNNNTAGAD 431	DR	DEGK BOX.	
ID	YAB9_YEAST	STANDARD:	Qy 3 DDGDDNDGGGENGGCDNYATDD 27	DR	DEGK BOX.	
AC	P11380;	PRT; 1131 AA.	Db 407 DDNDNDGNNNNNNNNNNNTAGAD 431	DR	DEGK BOX.	
DT	01-JUL-1993	(Rel. 26, Created)	RESULT 5	DR	DEGK BOX.	
DT	16-OCT-2001	(Rel. 40, Last sequence update)	VG18_HSVI1	DR	DEGK BOX.	
DE	Hypothetical	128.5 kDa helicase in AT5-TPD3 intergenic region.	AC VG18_HSVI1	DR	DEGK BOX.	
GN	YAB019W OR YAL001	OR FUN30.	STANDARD;	DR	DEGK BOX.	
OS	Saccharomyces cerevisiae (Baker's yeast).	GN00120;	PRT;	DR	DEGK BOX.	
OC	Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;	DT 01-DEC-1992 (Rel. 24, Created)	318 AA.	DR	DEGK BOX.	
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.	DT 01-DEC-1992 (Rel. 24, Last sequence update)		DR	DEGK BOX.	
OX		DT 01-DEC-1992 (Rel. 24, Last annotation update)		DE	DEGK BOX.	
RN	[1] - TAXID=4932;			DE	DEGK BOX.	
RP	SEQUENCE FROM N.A.			DE	DEGK BOX.	
RC	STRAIN=S288C / AB972;			DE	DEGK BOX.	
RX	MEDLINE-93J209532;	PubMed=8458570;		DE	DEGK BOX.	
RA	Quellete B.F.F.; Clark M.W.; Keng T.; Storms R.K.; Zhong W.W.;			DE	DEGK BOX.	
RA	Barton A.B.; Kaback D.B.;			DE	DEGK BOX.	
RA	"Molecular cloning of chromosome I DNA from <i>Saccharomyces cerevisiae</i> : analysis of the 32 kb region between the LTE1 and SP07 genes.";			DE	DEGK BOX.	
RT	"Sequencing of chromosome I from <i>Saccharomyces cerevisiae</i> : analysis of a 32 kb region between the LTE1 and SP07 genes.";			DE	DEGK BOX.	
RN	[2]			DE	DEGK BOX.	
RP	SEQUENCE FROM N.A.			DE	DEGK BOX.	
RC	STRAIN=S288C / AB972;			DE	DEGK BOX.	
RX	MEDLINE-94193531;	PubMed=8144453;		DE	DEGK BOX.	
RA	Barton A.B.; Kaback D.B.;			DE	DEGK BOX.	
RT	"Molecular cloning of chromosome I DNA from <i>Saccharomyces cerevisiae</i> : analysis of the genes in the FUN38-MAK16-SP07 region.";			DE	DEGK BOX.	
RL	J. Bacteriol. 176:1872-1880(1994).			DE	DEGK BOX.	
RN	[3]			DE	DEGK BOX.	
RP	SEQUENCE FROM N.A.			DE	DEGK BOX.	
RC	STRAIN=S288C / AB972;			DE	DEGK BOX.	
RX	MEDLINE-92221690;	PubMed=1561836;		DE	DEGK BOX.	
RA	Clark M.W.; Zhong W.W.; Keng T.; Storms R.K.; Barton A.B.;			DE	DEGK BOX.	
RA	Kaback D.B.; Bussey H.;			DE	DEGK BOX.	
RT	"Identification of a <i>Saccharomyces cerevisiae</i> homolog of the SNF2 transcriptional regulator in the DNA sequence of an 8.6 kb region in the LTE1-CFS1 interval on the left arm of chromosome I.";			DE	DEGK BOX.	
RT	Yeast 8:133-145(1992).			DE	DEGK BOX.	
RL				DE	DEGK BOX.	
CC	- FUNCTION: MAY PROCESS OF UV DAMAGE REPAIR.	FUNCTION AS A NEGATIVE REGULATORY ELEMENT IN THE	Query Match	DR	DEGK BOX.	
CC	- SUBCELLULAR LOCATION: Nuclear (Potential)	Best Local Similarity 54.2%; Pred. No. 0.57; Mismatches 5; Indels 2; Gaps 1;	Score 70; DB 1; Length 318;	DR	DEGK BOX.	
CC	- SIMILARITY: BELONGS TO THE SNP2/RAD54 HELICASE FAMILY.	DB 294 DGDDGDDGDDGDDGDDGDD 317	Score 70; DB 1; Length 318;	DR	DEGK BOX.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	RESULT 6	DR 36047 MW; A30DFDB57C9E61 CRC64;	DR	DEGK BOX.	
CC	EMBL; L05116; AAC04338.1; -	IE68 PRVKA STANDARD;	PRT; 364 AA.	DR	DEGK BOX.	
DR	S22266; S22266	ID IE68 PRVKA		AC P24877;	DEGK BOX.	
DR	SGD; SGD_0000017; FUN30.	DT 01-MAR-1992 (Rel. 21, Created)		DT 01-MAR-1992 (Rel. 21, Last sequence update)	DEGK BOX.	
DR	GO:0007001; P:chromosome organization and biogenesis (se... . . ; IGI.	DT 01-FEB-1994 (Rel. 28, Last annotation update)		DT 01-FEB-1994 (Rel. 28, Last annotation update)	DEGK BOX.	

DB	Immediate-early protein RSP40.
GN	Pseudorabies virus (strain Kaplan) (PRV).
OS	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Alphaherpesvirinae; Varicellovirus.
OC	NCBI_TaxID:33703;
[1]	SEQUENCE FROM N.A.
RP	MEDLINE=91037977; PubMed=2172457;
RX	Zhang G., Leader D.P.; "The structure of the pseudorabies virus genome at the end of the inverted repeat sequences proximal to the junction with the short unique region.", J. Gen. Virology, 71:2433-2441(1990).
RX	-! SIMILARITY: BELONGS TO A FAMILY THAT GROUP TOGETHER HSV-1 AND HSV-2 IE-68 (US1), EHV-1 65, ERV-4 (ORF4), PRV RSP40, AND VZV 63.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
EMBL	DR 272588; CAA86770_1; -
EMBL	DR 272587; CAA86769_1; -
EMBL	DR 272573; S64073;
DR	SCD: S0003034; YGL066W.
DR	GO: GO:000124; C:SAGA complex; IDA.
DR	GO: GO:0016573; P:histone acetylation; IPI.
KW	Hypothetical protein.
FT	DOMAIN 141 150 POLY-ASP.
FT	DOMAIN 153 162 POLY-ASP.
FT	DOMAIN 466 476 POLY-GLN.
FT	DOMAIN 517 526 POLY-GLN.
SQ	SEQUENCE 657 AA; 72878 MW; 4618E2B47824321D CRC64;
Query Match	Score 25.4%; Length 557;
Best Local Similarity	40.7%; Pred. No. 1.2;
Matches	11; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
RESULT 8	
PRP3_RAT	ID PRP3_RAT STANDARD; PRT; 206 AA.
PRP3_RAT	ID PRP3_RAT P04474; AC P04474;
PR	DT 13-AUG-1987 (Rel. 05, Created)
PR	DT 13-AUG-1987 (Rel. 05, Last sequence update)
PR	DT 30-MAY-2000 (Rel. 39, Last annotation update)
PR	OS Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX	NCBI_TaxID=10116;
RN	SEQUENCE FROM N.A.
RP	TISSUE_Parotid Gland;
RC	MEDLINE=84289443; PubMed=6547951;
RX	Ziemer M.A., Swain W.F., Rutter W.J., Clements S., Ann D.K., Carlson D.M.;
RA	"Novel multigene families encoding highly repetitive peptide sequences. Sequence analyses of rat and mouse proline-rich protein genes.", RT peptide homologies of rat and human proline-rich proteins.", RL J. Biol. Chem. 259:10475-10480(1984).
RN	SEQUENCE OF 1-23 FROM N.A.
RC	TISSUE_Parotid Gland;
RX	MEDLINE=86033793; PubMed=3840480;
RA	Clement S., Mehansho H., Carlson D.M.;
RT	"Novel nucleotide sequence analysis of a proline-rich protein cDNA and peptide homologues of rat and human proline-rich proteins.", RL J. Biol. Chem. 260:13471-13477(1985).
CC	-! FUNCTION: MAY PROTECT TEETH BY BINDING TO TANNINS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch);
CC	SEQUENCE FROM N.A.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=S-88C;
RX	MEDLINE=97137793; PubMed=923474;
RA	Piagger M., Brueckner M., Schaefer M., Mueller-Auer S.; "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae chromosome VII.", RT Yeast 13:1077-1090(1997).
RN	[2]
RP	SEQUENCE OF 115-657 FROM N.A.
RC	STRAIN=S-88C;
RX	MEDLINE=97137793; PubMed=923474;
RA	Fuermann M. de Montigny J., Potier S., Souciet J.-L.; "The characterization of two new clusters of duplicated genes suggests a 'Lego' organization of the yeast Saccharomyces cerevisiae chromosomes.", RT Yeast 13:1061-1069 (1997).
CC	-! SIMILARITY: BELONGS TO THE ATAXIN 7 FAMILY.
CC	SEQUENCE OF 115-657 FROM N.A.
DR	EMBL; K0247; AAC41494_1;
DR	EMBL; M11898; AAC41495_1;
DR	EMBL; K0247; AAC41494_1;

KW	Zinc-finger; Repressor; Activator; Nuclear protein;	DR	EMBL; X87611; CAA60934.1; -.
KW	Metal-binding; DNA-binding; Repeat; 3D structure.	DR	EMBL; Z49512; CAA89536.1; -.
FT	ASP/GLU-RICH (ACIDIC).	DR	PIR; S55200; S55200.
FT	GLY-RICH.	DR	SGD; S0003773; YJR012C.
FT	POLY-HIS.	KW	Hypothetical protein; Transmembrane.
FT	GLY/SER-RICH.	FT	TRANSMEM 17 37 POTENTIAL.
FT	C2H2-TYPE.	SEQUENCE	207 AA; FDC11E6EDSAFA90A CRC64;
FT	C2H2-TYPE.	Query Match	24.0%; Score 68.5 ; DB 1; Length 207;
FT	C2H2-TYPE.	Best Local Similarity	50.0%; Pred. No. 0.54;
FT	INVOLVED IN NUCLEAR MATRIX ASSOCIATION.	Matches	13; Conservative 4; Mismatches 6; Indels 3; Gaps 1;
FT	INVOLVED IN REPRESSION OF ACTIVATED	Qy	2 NDDGGDNDGGGBENNGGDNAYATDD 27
FT	TRANSCRIPTION.	Db	174 NDSDNDNDNNNDNNKDGND--DD 196
FT	INVOLVED IN MASKING TRANSACTIVATION	RESULT 11	
FT	DOMAIN.	ID	KEX2_CANAL STANDARD;
FT	H -> R (IN REF. 2).	ID	KEX2_CANAL AC 013359;
FT	G -> R (IN REF. 1).	DT	15-JUL-1998 (Rel. 36, Created)
FT	CONFLICT	FT	15-JUL-1998 (Rel. 36, Last sequence update)
FT	65	FT	16-OCT-2001 (Rel. 40, Last annotation update)
FT	CONFFLICT	DE	Kexin precursor (EC 3.4.21.61) (KEX2 protease).
FT	196	GN	KEX2
FT	196	OS	Candida albicans (Yeast).
FT	196	OC	Bukaryota; Fungi; Ascomycota; Saccharomycetes;
FT	STRAND	OC	Saccharomycetales; mitosporic Saccharomycetes; Candida.
FT	TURN	RN	[1]_NCBI_TaxID5476;
FT	TURN	RP	SEQUENCE FROM N.A.
FT	STRAND	RC	STRAIN=WO-1;
FT	STRAND	RX	Medline=98030572; PubMed=9360967;
FT	STRAND	RA	Newport G.R., Arabian N.;
FT	STRAND	RT	"KEX2 influences Candida albicans proteinase secretion and hyphal
FT	STRAND	RT	formation.";
FT	HELIX	RL	J. Biol. Chem. 272:28954-28961(1997);
FT	HELIX	CC	- - CATALYTIC ACTIVITY: Cleavage of Lys Arg-Xaa and Arg-Arg-L-Xaa
FT	HELIX	CC	bonds to process Yeast alpha-factor pheromone and killer toxin
FT	HELIX	CC	precursors.
FT	HELIX	CC	- - COFACTOR: CALCIUM DEPENDENT (BY SIMILARITY).
FT	HELIX	CC	- - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. LATE GOLGI
FT	HELIX	CC	COMPARTMENT (BY SIMILARITY).
FT	HELIX	CC	- - PTM: O-GLYCOSYLATED (BY SIMILARITY).
FT	HELIX	CC	- - SIMILARITY: BELONGS TO PEPTIDASE FAMILY 38. FURIN SUBFAMILY.
FT	HELIX	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	HELIX	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	HELIX	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	HELIX	CC	use by non-profit institutions as long as its content is in no way
FT	HELIX	CC	modified and this statement is not removed. Usage by and for commercial
FT	HELIX	CC	entities requires a license agreement (See http://www.ebi-sib.ch/announce/
FT	HELIX	CC	or send an email to license@ebi-sib.ch).
FT	HELIX	CC	DR EMBL; AF02372; AAB80929.1; -.
FT	HELIX	CC	DR HSSP; P00752; 28BT.
FT	HELIX	CC	DR MEROPS; S08_070; -.
FT	HELIX	CC	DR InterPro; IPR002884; P_domain.
FT	HELIX	CC	DR Pfam; PF01483; P_proprotein; PARTIAL.
FT	HELIX	CC	DR PFAM; PP00082; Peptidase_S8; 1.
FT	HELIX	CC	DR PRODOM; PD000723; SUBTILASE_S8.
FT	HELIX	CC	DR PROSITE; PS00136; SUBTILASE_AS; FALSE_NEG.
FT	HELIX	CC	DR PROSITE; PS00137; SUBTILASE_HIS; 1.
FT	HELIX	CC	DR PROSITE; PS00138; SUBTILASE_SRR; 1.
FT	HELIX	CC	KW Hydrolyse; Serine proteases; Glycoprotein; Calcium; Transmembrane;
FT	HELIX	CC	KW Zymogen; Signal; 1 20
FT	HELIX	CC	FT SIGNAL 1 20
FT	HELIX	CC	FT PROPEP 21 2
FT	HELIX	CC	FT CHAIN ? 924
FT	HELIX	CC	FT DOMAIN ? 774
FT	HELIX	CC	POTENTIAL.
FT	HELIX	CC	POTENTIAL.
FT	HELIX	CC	REXIN.
FT	HELIX	CC	LUMENAL (POTENTIAL).

DR PROSITE; PS00863; CALSEQUESTRIN_1; 1;	DR PRINTS; PF00182; Glyco_hydro_19; 1.
DR PROSITE; PS00864; CALSEQUESTRIN_2; 1;	DR PRODOM; PD000609; Chitin_binding_1; 1.
KW Muscle; Glycoprotein; Calcium-binding; Signal.	DR PRODOM; PD354900; Glyco_Hydro_19; 1.
FT SIGNAL 1 19	DR SMART; SM00270; ChitBD1_1.
FT CHAIN 20 409 CALSEQUESTRIN, CARDIAC MUSCLE ISOFORM.	DR PROSITE; PS00026; CHITIN_BINDING; 1.
FT DOMAIN 373 409 ASP/GLU-RICH (ACIDIC).	DR PROSITE; PS00073; CHITINASE_19_1; 1.
FT CARBOHYD 335 335 N-LINKED (GLCNAC). .) (POTENTIAL).	DR PROSITE; PS00074; CHITINASE_19_2; 1.
FT CONFLICT 45 45 I -> S (IN REF. 2).	KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding;
SQ SEQUENCE 409 AA; 47356 MW; 068D469EA7457DAE CRC64;	KW Multigene family; Signal; BY SIMILARITY.
Query Match Score 67; DB 1; Length 409;	FT SIGNAL 1 22 BY SIMILARITY.
Best Local Similarity 24.3%; Pred. No. 1.6;	FT CHAIN 2 3 340 ACIDIC ENDOCHITINASE WIN6.
Matches 12; Conservative 5; Mismatches 10; Indels 0;	FT DOMAIN 2 3 62 CHITIN-BINDING (BY SIMILARITY).
Qy 1 DNDDGDDNDGGGENNGGDDNYATDD 27	FT DOMAIN 6 3 85 SPACER.
Db 382 DDDDDDDDDNGNSDEEDDDSDDED 408	FT CHITINASE; BY SIMILARITY.
RESULT 14	FT DISULFID 25 40 BY SIMILARITY.
CHI6_POPTR STANDARD PRT; 340 AA.	FT DISULFID 3 4 46 BY SIMILARITY.
AC P16579; ID CHI6_POPTR STANDARD; PRT; 340 AA.	FT DISULFID 3 9 53 BY SIMILARITY.
AC P16579; ID CHI6_POPTR STANDARD; PRT; 340 AA.	FT DISULFID 5 7 61 BY SIMILARITY.
DT 01-OCT-1996 (Rel. 15, Created)	FT DISULFID 11 0 172 BY SIMILARITY.
DT 01-OCT-1996 (Rel. 34, Last sequence update)	FT DISULFID 18 3 191 BY SIMILARITY.
DB Acidic endochitinase WIN6 precursor (EC 3.2.1.14).	FT DISULFID 29 0 323 999 MW; C4E96473BEA55CS CRC64;
GN WIN6.	SQ 340 AA; 3640 AA; C4E96473BEA55CS CRC64;
OS Populus trichocarpa (Western balsam poplar).	Query Match Score 66.5; DB 1; Length 340;
OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina;	Best Local Similarity 56.0%; Pred. No. 1.5;
OC Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Rosidae;	Matches 14; Conservative 4; Mismatches 4; Indels 3; Gaps 2;
OC eurosids I; Malpighiales; Salicaceae; Populus.	SEQUENCE 5 GDDNDDG - GGENNGGDGN - YATD 26
OX NCBI_TaxID=3694;	67 GDGDDGCGDDGDDGDDGYLSD 91
RN [1]	RESULT 15
RP SEQUENCE FROM N.A. CAQ2_RAT	CAQ2_RAT
STRAIN=11-11; TISSUE=Leaf; RX ID CAQ2_RAT STANDARD; PRT; 413 AA.	ID CAQ2_RAT STANDARD; PRT; 413 AA.
MEDLINE=9435653; PubMed=805397; AC P51868; O09177;	AC P51868; O09177;
RA Clarke H.R.; Davis J.M.; Wibert S.M.; Bradshaw H.D. Jr., Gordon M.P.; DT 01-OCT-1996 (Rel. 34, Created)	DT 01-OCT-1996 (Rel. 34, Last sequence update)
RT "Wound-induced and development activation of a poplar tree	DT 28-FEB-2003 (Rel. 41, Last annotation update)
DB chitinase gene promoter in transgenic tobacco."; Plant Mol. Biol. 25:799-815(1994).	DB Calsequestrin, cardiac muscle isoform precursor (Calsequestrin 2).
RN [2]	GN CASQ2.
RP SEQUENCE OF 127-340 FROM N.A. OS Rattus norvegicus (Rat).	OS Rattus norvegicus (Rat).
RX MEDLINE=90046703; PubMed=2833366; OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Rattus.
RA Parsons T.J.; Bradshaw H.D. Jr., Gordon M.P.; RT Sequence of specific mRNAs in response to wounding in	RC STRAIN=Sprague-Dawley; TISSUE=Heart muscle;
RT "Systemic accumulation of specific mRNAs in response to wounding in	RA SPRAIN=Sprague-Dawley; TISSUE=Heart muscle;
RT poplar trees"; RT SEQUENCE OF 50-413 FROM N.A.	RA Aquilla T.T.; Rovner A.S.;
CC INDUCTION: By wounding.	RL Submitted (AUG-1995) to the EMBL/GenBank/DDBJ databases.
CC FUNCTION: THIS PROTEIN FUNCTIONS AS A DEFENSE AGAINST CHITIN CC SEQUENCE FROM N.A.	RN [1]
CC CONTAINING FUNGAL PATHOGENS.	RP STRAIN=Sprague-Dawley;
CC CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-	RC Rodriguez M.M.; Chen C.; Smith B.; Mochly-Rosen D.;
CC acetyl-D-glucosamine polymers of chitin.	RA Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC -I- INDUCTION: By wounding.	CC FUNCTION: CAUSEQUESTRIN IS A HIGH-CAPACITY, MODERATE AFFINITY,
CC -I- SIMILARITY: BELONGS TO CHITINASE CLASS IA WHICH IS SIMILAR TO	CC CALCIUM BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE
CC CLASS IB IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERMINAL	CC IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALEQUESTRIN THROUGH
CC CHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GLYCOSYL	CC A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40
CC HYDROLASES).	CC -I- SUBCELLULAR LOCATION: THIS ISOFORM OF CALCIUM (BY SIMILARITY).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	CC -I- SARCOPLASMIC RETICULUM'S TERMINAL LUMINAL SPACES OF
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	CC CARDIAC AND SLOW SKELETAL MUSCLE CELLS
CC the European Bioinformatics Institute. There are no restrictions on its	CC -I- SIMILARITY: BELONGS TO THE CALEQUESTRIN FAMILY.
CC use by non-profit institutions as long as its content is in no way	CC
CC modified and this statement is not removed. Usage by and for commercial	CC
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	CC
CC or send an email to license@isb-sib.ch).	CC
DR U01661; AAA57278_1; -	CC This SWISS-PROT entry is copyright. It is produced through a collaboration
DR U01660; AAA57277_1; -	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
DR M25336; AAA56701_1; -	CC the European Bioinformatics Institute. There are no restrictions on its
DR HSSP; P02877; 1HEV	CC use by non-profit institutions as long as its content is in no way
DR InterPro; IPR01002; Chitin binding_1	CC modified and this statement is not removed. Usage by and for commercial
DR InterPro; IPR000736; Glyco_Hydro_19_1.	CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
DR Pro0187; Chitin_bind_1; 1.	CC or send an email to license@isb-sib.ch).

CC
DR EMBL; U33287; AAA75480.1;
DR EMBL; AF001334; AAB58746.1;
DR HSSP; P07221; 1A8Y.
DR InterPro; IPR001393; Calsequestrin.
DR Pfam; PF01216; Calsequestrin; 1.
DR PRINTS; PR00312; CALSEQUESTRIN.
DR PROSITE; PS00863; CALSEQUESTRIN_1; 1.
DR PROSITE; PS00864; CAUSEQUBSTRIN_2; 1.
KW Muscle; Glycoprotein; Calcium-binding; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 413 CALSEQUESTRIN, CARDIAC MUSCLE ISOFORM.
FT DOMAIN 373 413 ASP/GLU-RICH (ACIDIC).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 50 52 DLL -> ARV (IN REF. 2).
FT CONFLICT 208 211 FLEV -> SUKM (IN REF. 2).
SQ SEQUENCE 413 AA; 47868 MW; 1ADEC534250D9C26 CRC64;
Query Match 24.1%; Score 66.5; DB 1; Length 413;
Best Local Similarity 51.9%; Pred. No. 1.0;
Matches 14; Conservative 3; Mismatches 9; Indels 1; Gaps 1;
Qy 1 DNDDGDDNDGGGENNGGDSNYADD 27
Db 387 DNDDDDDDNSDEDN-DDSDDDDD 412

Search completed: December 3, 2003, 16:18:43
Job time : 12 secs

	result	Query	Match	Length	ID	Description
No.	Score	No.	Score	No.	Score	No.
1:	sp_archaea:*	1:	99.5	36.1	1360	5 Q94649 plasmodium
2:	sp_bacteria:*	2:	92	33.3	1249	5 Q812R2 plasmodium
3:	sp_fungi:*	3:	85.5	31.0	1864	5 Q8icp5 plasmodium
4:	sp_human:*	4:	85	30.8	899	5 Q8ibnb8 plasmodium
5:	sp_invertebrates:*	5:	85	30.8	2961	5 Q8i452 plasmodium
6:	sp_mammal:*	6:	84.5	30.7	894	5 Q8i511 plasmodium
7:	sp_micr:*	7:	82	30.7	734	10 Q9lnz0 arabidopsis
8:	sp_organelle:*	8:	81.5	29.5	1176	5 Q8sbv4 dictyostel
9:	sp_phage:*	9:	81	29.3	1308	5 Q8ik00 plasmodium
10:	sp_plant:*	10:	80	29	90	4 Q9By99 homo sapien
11:	sp_rodent:*	11:	80	29.0	2276	2 Q93Ty6 staphylococ
12:	sp_virus:*	12:	78.5	28.4	1559	5 Q8i406 plasmodium
13:	sp_vertebrate:*	13:	78	28.3	447	12 Q9YTJ3 ateline h
14:	sp_unclassified:*	14:	78	28.3	966	5 Q8iLk5 plasmodium
15:	sp_virus:*	15:	77.5	28.1	4095	5 Q8icn0 plasmodium
16:	sp_bacteriap:*	16:	77	27.9	393	5 Q26592 schistosom
17:	sp_archeap:*					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

17	77	27.9	764	3	Q12500	saccharomyces
18	75	27.9	1946	5	Q97391	plasmodium
19	76.5	27.7	607	10	Q9rG45	arabidopsis
20	76.5	27.7	1105	5	Q8TM35	plasmodium
21	76.5	27.7	1431	5	Q81UH5	plasmodium
22	76	27.5	222	5	Q96629	plasmodium
23	76	27.5	461	12	Q93861	spodoptera
24	75.5	27.4	669	5	Q8SX8	drosophila
25	75.5	27.4	669	5	Q9VYB6	drosophila
26	75.5	27.4	1641	5	Q8IKZ0	plasmodium
27	75	27.4	5439	5	Q81586	plasmodium
28	75	27.2	382	1.0	Q9JK24	arabidopsis
29	75	27.2	945	5	Q8JY8	plasmodium
30	75	27.2	2162	5	Q8IBH2	plasmodium
31	75	27.2	2754	5	Q8I151	plasmodium
32	74.5	27.0	1371	5	Q8ICK4	plasmodium
33	74	26.8	189	5	Q9NHZ0	drosophila
34	74	26.8	420	5	Q8JJE1	plasmodium
35	74	26.8	1125	5	Q8BBU0	plasmodium
36	74	26.8	1173	5	Q8NS50	drosophila
37	74	26.8	5779	5	Q8IBS0	plasmodium
38	73.5	26.6	299	10	Q8H3D9	oryza sativa
39	73.5	26.6	2820	5	Q8IM32	plasmodium
40	73	26.4	297	5	Q8JHT0	plasmodium
41	73	26.4	366	5	Q815P9	plasmodium
42	73	26.4	411	11	Q8CHJ4	rattus norvegicus
43	73	26.4	541	10	Q9MA97	arabidopsis
44	73	26.4	682	5	Q8IJ98	plasmodium
45	73	26.4	780	5	Q8IL2Z	plasmodium

ALIGNMENTS

RESULT 1			
Q94649	PRELIMINARY;	PRT;	<u>1360 AA.</u>
ID			
Q94649;			
AC			
Q94649;			
DT	01-FEB-1997	(TREMBLrel.	02, Created)
DT	01-JAN-1998	(TREMBLrel.	05, Last Sequence update)
DT	01-MAR-2003	(TREMBLrel.	23, Last annotation update)
DE	AARP2.		
GN			
OS	Plasmodium falciparum (isolate 3D7).		
OC	Bukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI TaxID:61329.		

[1]	SEQUENCE FROM N.A. STRAIN=317;	SEQUENCE FROM N.A. STRAIN=317;	SEQUENCE FROM N.A. STRAIN=317;
MEDLINE=97391121; PubMed=9247928; Barale J.C., Attal-Bonnefoy G., Brahimi K., Pereira L., Langsley G.; "PfAAPR2 is a novel evolutionary conserved protein whose repeats identify a new antigen family in Plasmodium falciparum.";	Barale J.C., Candelle D., Attal-Bonnefoy G., Dehoux P., Bonnefoy S., Ridley R., Pereira da Silva L., Langsley G.; "Plasmodium falciparum AARPI, a giant protein containing repeated motifs rich in asparagine and aspartate residues, is associated with the infected erythrocyte membrane."; Infect. Immun. 65:3003-3010(1997). EMBL: Y08924; CAA10129.1; -. InterPro: IPR007034; DUF663. PFAM: PF04950; DUF663; 1.	SEQUENCE 1360 AA; 158109 MW; 400FFRAC9C74AE3A CRC64;	every Match 36.1%; Score 99.5%; DB 5; Length 1360; at Local Similarity 40.4%; Pred. No. 0.001; Indels 11; Gaps matches 23; Conservative 8; Mismatches 15;

RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.

RN [2] SEQUENCE FROM N.A. PubMed=12368867;
 RX MEDLINE=22255708; Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D., Bowman S., Atkin R., Baker S., Barron A., Brooks K., Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Christodoulou Z., Clark L., Clark R., Cotton C., Chillingworth T., Cronin A., Davies R., Dear P., Deardon F., Doggett J., Fellwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z., Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P., Knights A., Konforovic B., Kyes S., Jagels K., James K.D., Johnson D., Kerhornou A., Line A., Maddison M., McLean J., McLean J.J., Mooney P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabbinowitsch E., Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K., Sharp S., Smith P., Squares R., Stevens K., Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J., Sulston J.E., Craig A., Newbold C., Barrell B.G.; RT Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13. ";
 RL Nature 419:531(2002).
 DR AL929352; CD51463.1;
 KW Kinase; Transferase.
 SEQUENCE 1559 AA; MW: 9087140E1E176E96 CRC64;

Query Match Score 78.5%; DB 5; Length 1559;

Best Local Similarity 35.4%; Pred. No. 0.45%; Matches 17; Conservative 7; Mismatches 11; Indels 13; Gaps 1;

Qy 1 DNDDGGDDNDGGGENNGGGDN-----YATDDMMWMM 35
 Db 131. DDDDGDDDDDDGDDGDDDSNSNDIKNHKYNDYHOMIKL 178

RESULT 13
 Q9YTJ3 PRELIMINARY; PRT; 447 AA.
 AC Q9YTJ3; DT 01-MAY-1999 (TREMBrel. 10, Created)
 DT 01-MAY-1999 (TREMBrel. 10, Last sequence update)
 DT 01-MAR-2003 (TREMBrel. 23, Last annotation update)
 DE Orf 73.
 OS Atelina herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 NCBI_TaxID=85618;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=73;
 RX MEDLINE=20091363; PubMed=10623770;
 RA Albrecht J.C.; RT "Primary structure of the Herpesvirus Ateles genome.";
 RL J. Virol. 74:1033-1037(2000).
 RN [2] SEQUENCE FROM N.A.
 RP Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AF083424; AAC9598.1; -
 SEQUENCE 447 AA; MW: EA242FE09A6A5EB9 CRC64;

Query Match Score 78%; DB 12; Length 447;
 Best Local Similarity 65.0%; Pred. No. 0.14%; Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DNDDGGDDNDGGGENNGGG 20
 Db 176 DEDGDDDEDGSDGGDGD 195

RESULT 14
 Q8ILKS PRELIMINARY; PRT; 966 AA.
 AC Q8ILKS; DT 01-MAR-2003 (TREMBrel. 23, Created)
 DT 01-MAR-2003 (TREMBrel. 23, Last sequence update)
 DE Hypothetical protein.
 OS Plasmodium falciparum (isolate 3DP).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=3DP;
 RC MEDLINE=22255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Rung R., White O., Berriman M., Hyman R.W., Carter J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S., Chan M.-S., Nete V., Shallom S.J., Suh B., Peterson J., Angiuoli S., Perera M., Allen J., Selengut J., Haft D., Mather M.W., Roos D.S., Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A., McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C., Newbold C., Hoffman S.L., Newbold C., Davis R.W., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M., Barrell B.; RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
 RL Nature 419:498-511(2002).
 DR EMBL: AB014820; AAC36851.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 966 AA; MW: 77D45968C5250BB7 CRC64;

RESULT 15
 Q8ICNO PRELIMINARY; PRT; 4095 AA.
 AC Q8ICNO; DT 01-MAR-2003 (TREMBrel. 23, Created)
 DT 01-MAR-2003 (TREMBrel. 23, Last sequence update)
 DE Orf 73.
 OS Plasmodium falciparum (isolate 3DP).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 RN [1] SEQUENCE FROM N.A.
 RP Chernevach I., Davis P., Goodhead I., Stevens K., Mungall K., Berriman M., Pain A., Hall N., Chillingworth C., Doggett J., Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.; RT Submitted (SBP-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AL84405; CAD50411.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 4095 AA; MW: B64911B1716E0D12E CRC64;

Query Match Score 28.1%; DB 5; Length 4095;
 Best Local Similarity 46.7%; Pred. No. 1.6%; Matches 14; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 2 NDGGDDNDGGGENNGGG-GDNYATDDNNM 30
 Db 1908 DDNGDDNDGDDNGDDNGDDNGDDYDEDNIKL 1937

T42987 hypothetical protein 73 - ateline herpesvirus 3
 C;Species: ateline herpesvirus 3
 C;Variety: Strain 73
 C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C;Accession: T42987
 R;Albrecht, J.C.; Fleckenstein, B.
 submitted to the EMBL Data Library, August 1998
 A;Description: Primary structure of the herpesvirus atelies genome.
 A;Reference number: 222274
 A;Accession: T42987
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-447 <ALB>
 A;Cross-references: ENBL:At083444 ; PIDN:ARC95538.1
 A;Experimental source: strain 73
 C;Superfamily: herpesvirus immediate-early protein IE68
 Query Match, Score 78; DB 2; Length 447;
 Best Local Similarity 65.0%;
 Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 DNDDGDDNDGGGBENNGGDD 20
 Db 176 DGEDGDDDEDGGDGGDGGDGG 195

RESULT 4
 S64951 hypothetical protein YLR14c - yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein L2941
 C;Species: Saccharomyces cerevisiae
 C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
 C;Accession: S64951; S64901
 R;Verhasselt, P.; Voet, M.; Volckaert, G.
 submitted to the Protein Sequence Database, May 1996
 A;Reference number: S64943
 A;Accession: S64951
 A;Molecule type: DNA
 A;Residues: 1-764 <VER>
 A;Cross-references: ENBL:273286; PIDN:CAA97681.1; PID:e245807; PID:9136051
 A;Experimental source: strain S28C
 R;Verhasselt, P.; Volckaert, G.
 submitted to the EMBL Data Library, September 1995
 A;Reference number: S649393
 A;Accession: S64901
 A;Molecule type: DNA
 A;Residues: 1-764 <VER>
 A;Cross-references: ENBL:X89514; PIDN:CAA61692.1; PID:e198747; PID:9129702
 C;Genetics:

A;Gene: SGD:EFRA
 A;Cross-references: SGD:S0004104
 A;Map Position: 12R

Query Match, Score 77; DB 2; Length 764;
 Best Local Similarity 48.1%;
 Matches 13; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 DNDDGDDNDGGGENNNGGDNATDD 27
 Db 625 DSEDDDDDDGEGDENGDEGDEG 651

RESULT 5
 S22266 FUN30 protein - yeast (Saccharomyces cerevisiae)
 N;Alternate name: protein YAL001; protein YAL019w
 C;Species: Saccharomyces cerevisiae
 C;Accession: S22266; S36715
 R;Clark, M.W.; Zhong, W.W.; Keng, T.; Storms, R.K.; Barton, A.; Kaback, D.B.; Bussey, H.
 Yeast 8, 133-145, 1992
 A;Title: Identification of a Saccharomyces cerevisiae homolog of the SNF2 transcriptional

A;Reference number: S22266; MUID:92221690; PMID:1561836
 A;Accession: S22266
 A;Molecule type: DNA
 A;Residues: 1-1131 <CLAA>
 A;Cross-references: GB:L05146; EMBL:S93805; NID:9171851; PIDN: AAC04938.1; PID:9171856
 R;Ouellette, F.; Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Zeng, B.; Portin, N.;
 submitted to the EMBL Data Library, January 1993
 A;Description: Sequencing of Chromosome I from Saccharomyces cerevisiae: analysis of a
 A;Reference number: S36715
 A;Accession: S36715
 A;Molecule type: DNA
 A;Residues: 1-1131 <OUE>
 A;Cross-references: EMBL:L05146; NID:9171851; PIDN: AAC04938.1; PID:9171856; MIPS:YAL019
 C;Genetics:
 A;Gene: SGD:FDN30
 A;Cross-references: SGD:S000017; MIPS:YAL019w
 A;Map position: 1L
 C;Keywords: DNA binding; nucleus; transmembrane protein
 F;969-985;Domain: transmembrane #status predicted <TM1>
 Query Match, Score 73; DB 2; Length 1131;
 Best Local Similarity 52.0%;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 Qy 3 DGDDNDGGGENNNGGDNATDD 27
 Db 407 DDDDDNDGNNNDNNNNNNNTAGAD 431

RESULT 6
 P69069 hypothetical protein MTH1519 - Methanobacterium thermoautotrophicum (strain Delta H)
 C;Species: Methanobacterium thermoautotrophicum
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1999
 C;Accession: P69069
 R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadefora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noeling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
 A;Reference number: A69000; PMID:98037514; PMID:971463
 A;Accession: P69069
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-91 <MTH>
 A;Cross-references: GB:AE000912; GB:AE000666; NID:92622632; PIDN:AA885994.1; PID:926226
 A;Experimental source: strain Delta H
 C;Genetics:
 A;Gene: MTH1519

Query Match, Score 72; DB 2; Length 91;
 Best Local Similarity 29.5%;
 Matches 13; Conservative 16; Mismatches 11; Indels 4; Gaps 1;

Qy 4 DGDDNDGGGENNNGGDNATDDMMMMVVVMMVVMMVVI 47
 Db 50 DDDDDSGGADDSAGSDWL---WLVVLLIVVAVLILYL 89

RESULT 7
 T30613 hypothetical protein 11L - Molluscum contagiosum virus 1
 N;Alternate names: Mc011L
 C;Species: Molluscum contagiosum virus 1
 C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999
 C;Accession: T30613
 R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
 Science 273, 813-816, 1996
 A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host
 A;Reference number: 220876; MUID:96325459; PMID:8670425
 A;Accession: T30613
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA

A;Cross-references: EMBL:Z72588; NID:91322573; PIDN:CAA96770.1; PID:e2443288; PID:9132257				
A;Experimental source: strain S288C				
R.Feuermann, M.; Pötter, S.; Souquet, J.L.				
submitted to the Protein Sequence Database				
A;Reference number: S64044				
A;Accession: S64070				
A;Molecule type: DNA				
A;Residues: 115-657 <PFEU>				
A;Cross-references: EMBL:Z72588; MIPS:YGL066W				
A;Experimental source: strain S288C				
C;Genetics:				
A;Cross-references: SGD:S0003034				
A;Map position: 7L				
Query Match	25.4%	Score 70;	DB 2;	Length 657;
Best Local Similarity	40.7%	Pred. No. 2.3;		
Matches 11;	Conservative	Mismatches 7;	Indels 0;	Gaps 0;
Qy	1 DNDDGGDDNDGGGENNGGDSNDNYATDD 27			
Db	147 DNDNDDDDDADDDDDNSNGANYKND 173			
RESULT 13				
T18501	hypothetical protein C0760c - malaria parasite (Plasmodium falciparum)			
C;Species: Plasmodium falciparum				
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000				
C;Accession: T18501				
R;Lawson, D.; Bowman, S.; Barrell, B.				
submitted to the EMBL Data Library, August 1997				
A;Reference number: Z18935				
A;Accession: T18501				
A;Status: Preliminary; translated from GB/EMBL/DBJ				
A;Molecule type: DNA				
A;Residues: 1-3394 <LAW>				
C;Genetics:				
A;Map position: 3				
A;Note: C0760c				
Query Match	25.4%	Score 70;	DB 2;	Length 3394;
Best Local Similarity	44.4%	Pred. No. 11;		
Matches 12;	Conservative	Mismatches 5;	Indels 0;	Gaps 0;
Qy	1 DNDDGGDDNDGGGENNGGDSNDNYATDD 27			
Db	3010 DDDNNDDDDDDDDNNDDDDNNNNDD 3036			
RESULT 14				
S52246	transposable retroelement homolog ; perennial teosinte			
N;Alternate names: ZLR57 protein, long repetitive sequence protein ZLRS 7				
C;Species: Zea diploperennis (perennial teosinte)				
C;Accession: JC4611; SS2246 #sequence_revision 21-Jul-1995 #text_change 07-Dec-1999				
C;Accession: JC4611; SS2246				
R;Montfort, A.; Vicent, C.M.; Raz, R.; Puigdomenach, P.; Martinez-Izquierdo, J.A.				
A;Title: Molecular analysis of a putative transposable retroelement from the Zea genus w				
A;Reference number: JC4641; MUID:97021439; PMID:8867795				
A;Accession: JC4641				
A;Molecule type: DNA				
A;Residues: 1-156 <MON>				
A;Cross-references: EMBL:X82087; NID:9609287; PIDN:CAA57619.1; PID:9609288				
Query Match	25.2%	Score 69.5;	DB 2;	Length 156;
Best Local Similarity	54.2%	Pred. No. 0.65;		
Matches 13;	Conservative	Mismatches 4;	Indels 3;	Gaps 1;
F;123-153; Region: DNA binding #status predicted				
1 DNDDGGDDNDGGGENNGG - -DGD 21				